DOCUMENT RESUME

ED 040 215 TE 001 911

AUTHOR	Millett, Nancy C.
TITLE	On Snarls and Straighteners.
INSTITUTION	Illinois Association of Teachers of English, Urbana.
PUB DATE	Oct 69
NOTE	6p.
JOURNAL CIT	Illinois English Bulletin; v57 n1 p13-18 Oct 1969
EDRS PRICE	EDRS Price MF-$0.25 HC-$0.40
DESCRIPTORS	*Composition Skills (Literary), English Instruction, Evaluation Criteria, *Evaluation Methods, *Grading, Literary Styles, Standards, Student Evaluation, *Teacher Responsibility, Teaching Methods

ABSTRACT

Inconsistent grading can be frustrating for students and embarrassing for teachers. To make grades more intelligible and systematic, a standard evaluation form such as the Diederich Scale can be valuable. Such a scale increases the teacher's responsibility and forces him (1) to define the meaning of the terms on the scale; (2) to define standards from "failure" to "excellent" for each item; (3) to evaluate the paper thoroughly; (4) to state the evaluation criteria; (5) to permit the student to revise the paper in response to concrete suggestions; (6) to judge the revision by the same standards as the original; (7) to given tangible and understandable rewards for revision; (8) to rank the importance of the items on the scale; (9) to correlate teaching emphasis with evaluation criteria; and (10) to teach aspects of style. Teachers may want to adapt standard scales to their individual teaching situations and to further differentiate and define the criteria, but a clearly understood grading scale can clarify the rules of the composition game for both teachers and students and lead to better cooperation and understanding. (LH)

Illinois English Bulletin;
Vol. 57, No. 1,
October, 1969

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED040215

TE001911

On Snarls and Straighteners

By Nancy C. Millett

Assistant Professor, College of Education, Wichita State University
Wichita, Kansas

Grading a student's paper can be as frustrating as trying to straighten a kid's snarled line while your own line needs rigging and everywhere around you the fish are jumping. If the light is good, your nails long and your fingers nimble, you might unsnarl the mess. But long before you're finished you long for a snag-proof line, a lash-proof reel, or some other kid. The tangle nastily present, you consider more realistic alternatives. You can, of course, cut bait and quit. Or you can cut, rerig, and tell the kid to fish close to shore where he might, by accident, catch a little one while you yourself go for a lunker. But if you really like the kid, you go right on picking away at the snarls. If you're lucky, you find a procedure to use on them and on all those snarls certain to follow, and if you're smart, you teach that procedure to the kid so that both of you can fish in earnest.

If you're like me, you're a much smarter fisherman than a grader of papers. It took me only one afternoon to develop a system for clearing a snarled line, but it's taken me fifteen years and more than twenty thousand papers to find a system for grading a student's prose. Obviously, I stand condemned as a slow learner — conservatively calculated, five thousand hours slow. But I *have* learned.

At first, prompted by more degreed colleagues who felt condemned to teach Freshman English, I learned to cut and quit. I looked at the messes just long enough to find one knot — a fragment, a run-on, or a comma fault — and I slashed it with an "F." And then, prompted by my undegreed students condemned to *take* Freshman English, I learned to shorten line. They fished dutifully for minnows, which I dutifully threw back with a "C." No one was having much fun, because the big ones still out there were still jumping. So I learned to glory in the kid's occasional far-

flung, long-curving, deep-driving cast — and as for the rest, I picked away, with increasingly nimble fingers, at the snarls. The experience was a salutary exercise in forbearance, and I might even have achieved a happy Mr. Chips-dom. Except that I slowly learned what my students soon discovered — the grades I assigned were not so much signs of their performance as of my caprice. What galled them, as it galled me, was the variability of my rewards for comparable first casts or recasts. "Nice try. 'C.'" Or "B." Or "A." All of us wanted me to be more reliable.

Thanks to chance, I soon had the opportunity to work on the problem in one of the centers for NEA's English Composition Project, and three years later, my grading was, in fact, highly reliable. Worked out during that time was a record and reward system, one that made sense to me, to other teachers, to lay readers, to statisticians, to administrators, *and,* so help me, to students.

During the first year, we worked with the Diederich Scale, a simple scoring system based on seven elements or cluster of elements to which professional writers and editors, lawyers, businessmen, and college English teachers, social science teachers, and natural science teachers responded when they evaluated an essay.[1] Urged by public school teachers to do so, Mr. Diederich added one more cluster: manuscript form and legibility. In its original form, the Diederich Scale coupled these eight items with a five-point scale, with "5" representing excellence. By circling one number in line with each item, a grader could record his evaluation of a

THE DIEDERICH SCALE

1 — Weak	2 — Fair	3 — Good	4 — Better	5 — Best
Quality and Development of Ideas			1 2 3 4 5	
Organization, Relevance, Movement			1 2 3 4 5	
Style, Flavor, Individuality			1 2 3 4 5	
Wording and Phrasing			1 2 3 4 5	
Grammar, Sentence Structure			1 2 3 4 5	
Punctuation			1 2 3 4 5	
Spelling			1 2 3 4 5	
Handwriting, Neatness			1 2 3 4 5	

Sum of Ratings _____

[1] Paul B. Diederich, John W. French, and Sydell T. Carlton, *Factors in Judgments of Writing Ability.* Research Bulletin 61-15. Princeton, N.J.: Educational Testing Service, 1961.

student's performance in each area. The sum of these scores represented overall performance, and that sum could, if the grader wished, be translated into a percent and/or into a letter grade. After grading a paper, the grader attached the marked scale to the student's theme.

Faulty though this scale proved to be, it was enormously helpful both to the grader and to the student. Attached to each student's paper, the scale encouraged talk about the worth of a paper in specific terms — terms more often embarrassing to the grader than to the student. Once the student was forced to see where his performance was weak and where it was strong, he began to ask why. His asking prodded even the most reluctant grader into a number of uncomfortable positions:

1. It forced the grader to define what he meant by each of the words on the scale and it forced him to explain those definitions to the student.
2. It forced the grader to define standards of failure, mediocrity, and excellence for each scaled item, and it forced him to state those standards to the student.
3. It forced the grader to do a thorough job of grading: it forced him to evaluate *every* element listed on the scale, it forced him to account for each judgment by offering specific marginal and summary comments, and it forced him to do so on each and every student's paper.
4. It forced the grader to state, *before* he read a set of papers, and even *before* he required his students to write, the criteria by which he would judge the student's performance.
5. It forced the grader to permit the student to revise his paper because, for the first time, the student could actually see that revision was in his own best interest.
6. It forced the grader to judge a revision by the *same* standards he had applied to the student's original draft.
7. It forced the grader to give tangible rewards for careful revision. If the student changed what the grader told him to change, his score *had* to go up — no matter how much the grader regretted, at second look, his first series of judgments.

In short, the Diederich Scale equipped the student to ask his teacher and his lay reader some wonderfully searching questions. For a while, we graders were mighty uncomfortable. But our troubles had scarcely begun. While we were groping for — and beginning to find — real answers to real demands, the brighter

student was raising sand on at least three other fronts. During the second and third years of the Project, his asking forced us to change the scale — and with it, our teaching:

8. It forced the grader to rank the items on the scale, to differentiate (as the scale clearly did not) between the importance of Ideas and Development, on the one hand, and Spelling or Punctuation or Handwriting on the other. That is, it forced the grader to decide *what* was important, *for whom* it was important, *when* and *why* it was that important, and just *how* important it was.
9. It forced the grader to bring his teaching of composition into line with the things that his use of the scale said he valued. If Quality and Development of Ideas counted for half a student's score, then a student's logic forced his grader to spend half his class time on modes of development rather than, say, on sentence drills.
10. It forced the grader to do what almost none of us had ever attempted to do — to teach aspects of style. Wording and Phrasing, and Style, Flavor, Individuality were items on the scale, but they were not in the curriculum. The students' questions forced us to teach what we said we could grade. Our discomfort was acute.

When we revised the Diederich Scale, we changed some of its wording and we also changed it to accommodate differences in grade level and ability level. For example, we grouped items on the scale, weighted each group, and weighted the groups differently for each grade level. A chart which summarizes these changes shows how simple they were to make, and how easy it would be to change weights simply by changing the multiplier for each group score:

ADAPTATIONS OF THE DIEDERICH SCALE

		F D C B A		WEIGHTS GRADE		
				7-8-9	10-11	12
CONTENT AND ORGANIZATION	Quality and Development of Ideas	1 2 3 4 5				
	Organization, Relevance, Movement	1 2 3 4 5	Sum × 5 ____	50%	50%	50%
STYLE	Wording and Phrasing	1 2 3 4 5	Sum × 1	10%		
	Style, Flavor, Individuality	1 2 3 4 5	or × 2 ____ or × 3		20%	30%
MECHANICS	Grammar, Sentence Structure	1 2 3 4 5				
	Punctuation	1 2 3 4 5	Sum × 2	40%		
	Spelling	1 2 3 4 5	or × 1.5 ____		30%	
	Manuscript Form, Legibility	1 2 3 4 5	or × 1			20%

Total Score (*Not* a Percent!) ____

No matter what weights we assigned to each area, we could easily convert the total scores into letter grades by consulting a simple conversion table:

CONVERSION TABLE

20-29	F
30-34	D−
35-44	D
45-49	D+
50-54	C−
55-64	C
65-69	C+
70-74	B−
75-84	B
85-89	B+
90-100	A

Thus adjusted, the scale apparently satisfied almost everyone. Other teachers and their lay readers were pleased for the same reasons that pleased me. I was happy because I finally had a system for recording the ways I judged a paper, a system that greatly reduced my caprice as I moved from paper to paper in a set, and from one set to another in a semester or a year. Because the scale forced me to judge the same elements for each and every student's paper, and because it forced me to give the same weight to those elements on each and every paper. I became a highly reliable grader. Especially improved was my grading of revised papers. All I had to do was look at the score I had assigned to each element, compare the student's revision with his original, and choose a higher (or, rarely, a lower) score. The revised scores, marked in blue next to the original scores marked in red, provided a graphic record of reward for my careful criticisms and the student's careful revisions. Together we had straightened the snarls; together we could go on fishing. Obviously, the student was happy too. At long last, he had forced his teacher to explain the rules of the game he was required to play, and at long last he could read the scoreboard — and change it, *his* way.

Curiously enough, even the administrators were pleased — not because they really knew what graders and students were doing for and with each other, but because their statisticians did know. Scores, after all, can be manipulated and used to report "results," expressed in statistical terms are what administrators read best.

More, much more, remains to be learned, and I am still learning, still slowly. The problem *now,* it seems to me, is to find terms that distinguish elements on the scale more precisely. "Wording and Phrasing," as distinct from "Style, Flavor, Individuality" is no distinction at all, or at best a sloppy one, isn't it? And I certainly am not happy to judge "Ideas and Development" separately from "Organization, Relevance, Movement." Some other classification, more nearly reflecting the connections between writer, subject, and audience — and between intention, structure, and effect — would be enormously helpful. I have some notions about this, and perhaps three years from now I will have picked my way through these new twists to old snarls. Meanwhile, a whole bunch of kids, and I, will go on fishing.